



TI 2007-008/3

Tinbergen Institute Discussion Paper

Growing richer and taller: Explaining Change in the Distribution of Child Nutritional Status during Vietnam's Economic Boom

Owen O'Donnell^{1,3}

Ángel López Nicolás²

Eddy van Doorslaer^{3,4}

¹ University of Macedonia, Greece;

² Universidad Politécnica de Cartagena, and Universitat Pompeu Fabra, Spain;

³ Erasmus University Rotterdam;

⁴ Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Roetersstraat 31

1018 WB Amsterdam

The Netherlands

Tel.: +31(0)20 551 3500

Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50

3062 PA Rotterdam

The Netherlands

Tel.: +31(0)10 408 8900

Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>.

Growing richer and taller: explaining change in the distribution of child nutritional status during Vietnam's economic boom

| | |
|---------------------|--|
| Owen O'Donnell | University of Macedonia, Greece & Erasmus University, the Netherlands |
| Ángel López Nicolás | Universidad Politécnica de Cartagena & Universitat Pompeu Fabra, Spain |
| Eddy Van Doorslaer | Erasmus University, the Netherlands |

27 November 2006

Abstract

Over a five-year period in the 1990s Vietnam experienced annual economic growth of more than 8% and a decrease of 15 points in the proportion of children chronically malnourished (stunted). We estimate the extent to which changes in the distribution of child nutritional status can be explained by changes in the level and distribution of income, and of other covariates. This is done using data from the 1993 and 1998 Vietnam Living Standards Surveys and a flexible decomposition technique that explains change throughout the complete distribution of child height. One-half of the decrease in the proportion of children stunted is explained by changes in the distributions of covariates and 35% is explained by change in the distribution of income. Covariates, including income, explain less of the decrease in very severe malnutrition, which is largely attributable to change in the conditional distribution of child height.

JEL Codes: I12, I31, O53

Keywords: Malnutrition, child height, decomposition, quantile regression, Vietnam

Correspondence: Owen O'Donnell, Department of Balkan, Slavic and Oriental Studies, University of Macedonia, 156 Egnatia Street, Thessaloniki 54006, Greece. Tel.: +30 2310 891478. Fax: +30 2310 891436. email: ood@uom.gr

Growing richer and taller: explaining change in the distribution of child nutritional status during Vietnam's economic boom

1 Introduction

Over a five-year period in the 1990s Vietnam experienced annual economic growth of more than 8% and a decrease of 15 points in the proportion of children suffering chronic malnourishment. Setting aside the econometrically challenging task of identifying the causal impact of income growth on the rate of malnutrition, it remains important to establish the extent to which changes in child nutritional status can be explained by, and are therefore predictable from, income growth. We provide this evidence using data from two waves of a Vietnamese survey and a flexible decomposition technique that makes it possible to explain change throughout the complete distribution of child nutritional status measured by height-for-age.

Reducing child malnutrition is a central goal of development (World Bank 2006). The first Millennium Development Goal sets a target of halving the proportion of undernourished children between 1990 and 2015. Predicting future rates of child malnutrition is important for the monitoring of development in general, and of progress toward the MDG target in particular. Predictions may be made from point-in-time relationships between some indicator of child nutritional status and a set of covariates, together with assumed changes in these covariates. There is particular interest in the decline in child malnutrition rates that can be expected to follow from assumed trends in economic growth (Haddad, Alderman, Appleton et al. 2003; Edmonds 2004). If declines in child malnutrition can be largely explained by rates of economic growth, then there is less need to monitor nutrition in addition to growth. It might even be argued that development targets defined in terms of nutrition become superfluous to those for economic growth.

The accuracy of predictions and the extent to which the evolution of child malnutrition can be traced from economic growth and changes in other determinants both depend upon whether there is stability in child nutrition functions over time. Relationships will not be stable if there are changes in the nature of the economic and public health environments that condition the determination of child nutrition. For

example, liberalisation of food and health care markets and the adoption of programmes specifically designed to correct nutrition deficiencies. Part of any change in the distribution of nutrition will be explained by changes in covariates and part by change in the nutrition function.

In this paper, we explain change in an indicator of long-term child nutritional status—height-for-age z-score¹—in Vietnam between 1993 and 1998. The experience of Vietnam in the 1990s provides an interesting case study of the evolution of child nutritional status during a period of rapid economic growth, transition to a market economy, nutrition policy initiatives and health sector changes. In the pre-transition era, child malnutrition remained high in Vietnam despite substantial falls in child mortality and fertility to levels significantly below those found in comparably poor countries (World Bank 2001). Part of the explanation was the inadequacy of energy and protein intake through a diet dominated by rice and severely lacking in protein from meat, pulses and fish (Hop 2003). The nutritional status of children improved substantially between 1993 and 1998, with the mean height-for-age z-score of children less than 10 years old increasing by 20% and the proportion of chronically undernourished, or stunted², children falling by 15 percentage points but remaining high at 37%. Inequality in nutritional status increased, with the coefficient of variation in height-for-age z-scores rising by 20% between 1993 and 1998. Over the same period, real GDP growth averaged 8.4% per annum, the poverty rate declined from 58% to 37%, the Gini coefficient of income inequality increased from 0.33 to 0.35 and markets continued to be liberalised and opened up to foreign competition (Glewwe 2003).

During the nineties there was a marked increase in efforts to tackle nutritional deficiencies in Vietnam. The Hunger Eradication and Poverty Reduction Program was introduced in 1992 and the first National Action Plan for Nutrition ran from 1995 to 2000 (Hop 2003). The latter set nutritional targets and was monitored and implemented by a newly established National Institute for Nutrition. A separate plan of action for iodine deficiency disorders was adopted in 1992. In agriculture the Government encouraged diversification and increased production of animal source foods with positive results (Hop 2003). Comparison of data from the 1987 and 2000 Nutrition Surveys reveals significant improvements in the quantity and quality of the diet (Hop 2003).

The health sector, which experienced severe financial pressure in the first period of economic reform from the late 1980s to early 1990s, fared much better over the 1993-98 period. Public health spending per capita increased by 50% in real terms, total spending on health increased from 5.2% to 8% of GDP, the real price of drugs fell by 30% due to deregulation of the pharmaceutical industry and liberalisation of the retail sector, commune level health services received much needed funds through government financing of health workers salaries and donor funding of health and family planning programmes (World Bank 2001). There was a 15% rise in real spending on disease control programmes, increased rates of child immunisation and falling prevalence of infectious diseases such as diarrhoea and, particularly, malaria (World Bank 2001). Social health insurance was introduced in 1993. The number of voluntary enrollees, the vast majority of which were school children, increased from 0.5 million in 1993 to 4 million in 1998 (World Bank 2001).

Economic growth is expected to improve child nutritional status by increasing households' means to purchase food and health care. But the nutrition policy initiatives and health sector changes implemented in Vietnam during the 1990s would also be anticipated to have positive consequences for child nutritional status. Some of these changes, such as increased immunisation and salt iodisation, may raise child nutrition for given values of observable individual and household level characteristics. They shift the intercept of a regression model. Others may change the way in which household characteristics are related to child nutrition. They change the slope coefficients of a regression model. For example, changes in the relative prices of food and medical care, will affect the relationship between household income and nutrition.

Application of the Machado and Mata (2005) decomposition method allows us to explain change throughout the full marginal distribution of height-for-age z-scores, distinguishing between the contributions of changes in the distributions of covariates, on the one hand, and those of changes in the parameters of the height-for-age function, on the other.³ Decomposition of change in the full distribution is important since severe malnourishment is of greater concern than mild malnourishment and factors contributing to changes at the bottom of the nutrition distribution need not be the same as those that explain change at the middle or top of the distribution. The approach allows us to explain change in the proportion of stunted children, which is of immediate interest, without imposing any distributional assumptions or assuming that the relationships of child height to covariates are constant across the conditional

distribution. Besides estimating the change in malnutrition that can be explained by changes in the distribution of all covariates, we also estimate the total and the partial contributions of changes in income. The latter is realised by an extension of the Machado and Mata method that exploits the panel nature of the data. We also explain the observed increases in inequality and the income gradient in child height deficit.

The remainder of the paper is organised as follows. In the next section, we place this paper within the context of the previous literature on the contribution of income growth to explanation of changes in child malnutrition. In the third section, we present the decomposition method. In section 4, we describe the dataset and present some descriptives. In section 5, we present the decomposition results. The final section concludes with an interpretation of the results.

2 Previous literature

The positive correlation between stature and income has long been recognised and has spurred a sizable literature on height as an indicator of living standards (Steckel 1995). Within the development economics literature, there has been considerable interest in recent years in explaining, and predicting, changes in child malnutrition from actual, and forecast, changes in income. (Smith and Haddad 2002) estimate that increases in GDP account for one-half of the fall in the rate on undernutrition—measured by weight-for-age more than two standard deviation below the median in the reference population—across 63 countries between 1970 and 1995. This income effect is estimated to operate through food availability, women’s education and status, sanitation and safe drinking water. Haddad, Alderman, Appleton et al. (2003) use both aggregate cross-country and household survey data to estimate the relationship between child nutrition status, measured by weight-for-age, and income, and conclude from the magnitude of the estimates obtained that even optimistic rates of income growth would not be sufficient to reach the MDG target on child malnutrition. The authors interpret their results as suggesting that economic growth alone cannot be relied upon to realise the target and programmes specifically targeted on nutrition are required.

The coincidence of high economic growth and a steep decline in child malnutrition in Vietnam in the 1990s has provoked research into the extent to which the latter is explained, if not caused, by the former. Ponce, Gertler and Glewwe (1998)

use the same general approach as Haddad, Alderman, Appleton et al. (2003) to predict the decline in child malnutrition rates in Vietnam that could be expected on the basis of forecast economic growth. From 1993 data, they estimate the relationship between height-for-age z-scores, household consumption per capita, used as a measure of income, and other covariates. Change in the proportion of children who are stunted is predicted by applying forecast rates of income growth to the estimated coefficient on consumption and adding the resultant scalar to the z-score of all children. This assumes not only inequality neutral growth of income and no change in other covariates, but also that the relationship between child height and income is constant across the conditional distribution of the former. The results suggest that reducing the proportion of stunted children under the age of five by 10 percentage points would require 6% growth in incomes per capita sustained for a period of 20 years. In fact, within five years at a growth rate very close to that forecast, the rate of stunting in this age group had fallen by 15 percentage points (Glewwe et al. 2003). This suggests factors other increasing incomes were mainly responsible for the decline in malnutrition. Glewwe, Koch and Nguyen (2003) and Edmonds (2004) both use data from the 1993 and 1998 Vietnam Living Standards Surveys (VLSS), as we do in this paper, to estimate the proportion of the actual change in height-for-age z-scores over the 5 year period that can be explained by increases in household income. Using the same approach as Ponce, Gertler and Glewwe (1998) to predict the incidence of stunting from cross-section estimates of the conditional mean z-score—consumption relationship, Glewwe, Koch and Nguyen (2003) find that consumption growth can explain at most 3.2 of the 15 percentage point fall in the proportion of rural children less than five years old who are stunted. Not only is the magnitude of the effect limited, it is based on an insignificant coefficient estimated by two-stage least squares.

If these estimates are correct and income growth does not account for most of the dramatic decline in child malnutrition in Vietnam in the 1990s, what does? Glewwe, Koch and Nguyen (2003) explore the possibility that increases in the quantity and quality of health services are largely responsible. They find only limited evidence in support of the hypothesis. Most of the available indicators of community level health services either show no substantial improvement over time, are not significantly correlated with height-for-age z-scores or have a only a small effect. Two exceptions are the existence of a sanitary toilet and a supply of oral rehydration

salts at the commune health centre. But it seems unlikely that these indicators explain the observed increase in nutrition.⁴

Edmonds (2004) seeks to establish the contribution of economic growth to the substantial improvement in child height in Vietnam in the 1990s while taking account of changes in the policy environment that condition the determination of child nutrition, and may be correlated with income changes. An instrumental variables (IV) solution is rejected given the difficulty of finding valid, non-weak instruments. This seems justified given the imprecision of the IV estimates obtained by Glewwe, Koch and Nguyen (2003). Instead, it is argued that since the policy environment is fixed at a given point in time, the cross-section correlation between height and income provides an appropriate estimate that can be used to determine the contribution of the observed change in income to that in height. This leads to an Blinder-Oaxaca type decomposition in which the change in mean height is explained by the change in income, on the one hand, and the change in the conditional mean function, on the other (Blinder 1973; Oaxaca 1973). Non-parametric regression is used to estimate a height-income relationship that is fully flexible over the income distribution. Unlike in Ponce, Gertler and Glewwe (1998) and Glewwe, Koch and Nguyen (2003), this avoids imposing the assumption that the income effect on height is the same irrespective of the level of income. Using this method, Edmonds (2004) finds that income growth can explain 60% of the 20% increase in the mean of height-for-age z-scores of children less than 10 years old. This is substantially more than the 20% of the 25% increase in the mean z-score of children less than 5 years attributed to income growth by Glewwe, Koch and Nguyen (2003). One explanation for the difference is that Edmonds' estimate is derived from the bivariate relationship between height and income and so reflects the contribution of income and all its correlates. Glewwe et al estimate the contribution of income change with all covariates held constant.

By allowing for non-linearity in the height-income relationship, Edmonds reveals that the contribution of income to height growth is substantially greater at lower initial levels of income. The restriction of this approach is that it estimates the income effect only on the conditional mean of height and decomposes change only in this parameter. It does explain change in other interesting features of the height-for-age distribution, such as the fall in the proportion stunted and the rise in inequality. The change in stunting could only be explained using this approach if one were to

assume that the conditional mean effect provides an adequate approximation to the association at other points in the conditional distribution. We relax this assumption by using quantile regression, allowing coefficients to vary across the conditional distribution of child height, and decomposing change in the full marginal distribution using the method of (Machado and Mata 2005). Unlike the Edmonds' approach, which is bivariate, we estimate a multivariate model. We estimate both the total and partial contribution of income to explanation of the change in the distribution of height. The total contribution is obtained, as in Machado and Mata, by changing the distribution of income and allowing all observable correlates to change with it. To get the partial contribution, we extend the Machado and Mata approach by exploiting the panel nature of the Vietnam Living Standards Surveys to change the distribution of income while keeping observable correlates constant.

Some of the studies referred to above attempted to identify the causal effect of income on child nutritional status by instrumenting income. Our goal is not to estimate causal effects but simply to explain the change in the child height distribution by separating the contribution of change in the distribution of covariates, including income, from that of change in their relationships with height. This allows us to establish the extent to which changes in child nutritional status are predictable from changes in covariates and track income growth. We cannot comment, however, on how child nutrition may change in response to an intervention that raised household incomes and kept all else constant.

3 Decomposition method

The (Machado and Mata 2005) method generalises the well-known (Blinder 1973)—Oaxaca (1973) decomposition to explain change in the full marginal distribution. To achieve this, one must obtain estimates that fully characterise the conditional distribution of the variable of interest and then simulate changes in the marginal distribution by combining these estimates with different samples of covariates. Full characterisation of the conditional distribution is achieved using quantile regression, allowing parameter heterogeneity across the conditional distribution. This is appropriate in the present context since we are particularly interested in changes occurring at low levels of nutrition and do not wish to assume that these can be predicted from conditional mean effects.

Simulations of the marginal and counterfactual distributions used in the decomposition are obtained as follows (Machado and Mata 2005). First, for each time period (t), quantile regression coefficients are estimated at a random sample of m quantile points. Second, for each t , a random sample of m sets of covariates is selected from the survey observations. We set m to 4500, which is slightly less than the number of children less than 10 years old in each of the 1993 and 1998 Vietnam Living Standards Surveys. A random sample of size m from the marginal distribution at time t is given by combining the randomly sampled covariates at t with the quantile regression coefficients estimated at t . Counterfactual distributions are constructed by combining the regression coefficients with the sampled covariates from the other time period.

Denote height-for-age z-score by h and a vector of covariates by X . Let $f(h_t)$ be the empirical density of h at t that is estimated directly from the data. Let $f^*(h_t)$ be the simulated marginal density constructed, as described in the previous paragraph, to be consistent with the quantile regression estimated conditional density. The estimate of the marginal density that would have prevailed in 1993 ($t=93$) if covariates had been distributed as in 1998 ($t=98$) is denoted by, $f^*(h_{93}; X_{98})$. Then, the change in any summary statistic of the distribution, $\alpha(\cdot)$, can be decomposed as follows:

$$(1) \quad \alpha(f(h_{98})) - \alpha(f(h_{93})) = \underbrace{\alpha(f^*(h_{98})) - \alpha(f^*(h_{93}; X_{98}))}_{\text{coefficients}} + \underbrace{\alpha(f^*(h_{93}; X_{98})) - \alpha(f^*(h_{93}))}_{\text{covariates}} + \text{residual}$$

The first and second terms on the right-hand side are the contributions of change in the coefficients and covariates respectively. The residual is due to the differences between the empirical densities and the marginal densities simulated from the quantile regression coefficients. As with all such methods, the decomposition is based on a particular weighting scheme. In (1), the coefficient contribution is weighted by period 1998 values of covariates, and the covariate contribution by period 1993 values of the coefficients. We check the sensitivity of our results to reversing the weights.

Equation (1) decomposes change in some parameter of the distribution into the contribution of changes in all coefficients and in all covariates. (Machado and Mata 2005) also suggest how to identify the contribution of a specific covariate by re-sampling from the simulated 1993 height density, $f^*(h_{93})$, such that the covariate is distributed as in 1998. Consider doing this for income. Percentiles of the 1998 income distribution are estimated, q_{98}^j , $j=1, \dots, J$. In this application, we use 20 evenly spaced percentiles. From $f^*(h_{93})$, select observations with income less than the first of these percentiles, q_{98}^1 , and re-sample from these with replacement such that their proportion in the sample corresponds to that in the 1998 distribution, i.e. 0.05 in this case. Repeat for all income ranges corresponding to the 1998 percentiles. From the resultant sample, one obtains an estimate of the marginal density of height z-scores under the counterfactual that income is distributed as in 1998 but the conditional density remains as in 1993. Denote this counterfactual density with $f^*(h_{93}; y_{98})$, where y indicates income. Then the estimated contribution of the change in the income distribution to the change in marginal distribution of height is given by,

$$(2) \quad f^*(h_{93}) - f^*(h_{93}; y_{98}) .$$

Note that this procedure holds the correlations of income with other covariates constant but not the values of those other covariates. As income is allowed to change, so will correlated covariates. It shows what would happen to child height if poor children in 1993 became like better-off children in 1993 in all observable characteristics. Not only would their income change, so would the education and height of their parents, and even their ethnicity. In this sense, the procedure is somewhat similar to the bivariate method of Edmonds (2004).

In order to estimate the partial contribution of income, we make use of the panel nature of the data. In the sample from which the 1993 height density is simulated, $f^*(h_{93})$, we replace each child's 1993 value of household income with the 1998 value of income for the same household and keep all other covariates (Z) at their 1993 values. This gives a counterfactual density, $f^*(h_{93}; Z_{93}, y_{98})$, representing how child height would have been distributed if it were determined as in 1993, all

covariates were as in 1993 but income was distributed as in 1998. The partial contribution of income is then given by,

$$(3) \quad f^*(h_{93}) - f^*(h_{93}; Z_{93}, y_{98}) .$$

The same approach is used to estimate the partial contribution of other covariates or sets of covariates.

Bootstrap confidence intervals are calculated for the estimated contributions of changes in covariates and in coefficients. This involves taking repeated samples from the original sample, taking account of the stratified and cluster sample design in doing so, and combining them with the quantile regression coefficients to obtain multiple estimates of the simulated marginal and counterfactual distributions.⁵

4 Data and descriptives

The Vietnam Living Standards Surveys provide data on 4799 households in 1992-93 and 5999 households in 1997-98. They form a panel with 4710 of the original households re-interviewed in 1997-98. Sampling was stratified by region and clustered by commune. The 1993 sample is self-weighted to be nationally representative. Supplementation of the sample in certain regions in 1998 introduced some non-representativeness that is corrected through the application of weights. The survey contains detailed measures of household expenditure and consumption. In common with standard practice, we use household consumption per capita as the indicator of household economic status and a proxy for permanent income. Child height was measured and transformed to a z-score value using the World Health Organisation (WHO) recommended US National Center for Health Statistics sample as the reference (World Health Organisation 1995).

We use the child as the unit of observation and, in each year, restrict attention to children under the age of 10 years. There are 5864 children in this age group in the 1993 sample and 5324 in 1998. We restrict the 1993 sample to children in households that had not dropped out of the survey by 1998. This is necessary since, as explained in the previous section, we use changes in household values to estimate the partial contribution of specific covariates. It results in the loss of 536 observations. We correct for any induced bias by applying inverse probability weights, which are calculated from probit estimates of the probability of household non-attrition as a

function of all covariates and the child's height-for-age z-score. Note that we do not require that the child is present in the household and under the age of 10 in both waves of the survey.

After the loss of observations missing on any of the covariates used in the quantile regressions, the sample sizes are 4895 and 4943 in 1993 and 1998 respectively. These are the observations from which samples are drawn to simulate the marginal and counterfactual densities. Following WHO recommendations on z-score values considered improbable and likely deriving from measurement error (World Health Organisation 1995), we set values above 3 and below -6 (i.e. 4 units below the 1993 mean) to missing. This results in 4794 observations in 1993 and 4852 in 1998 with valid z-score values and complete data on all covariates. These are the observations used in estimation of the quantile regressions.

FIGURE 1

The empirical density functions of height-for-age z-scores in 1993 and 1998 are presented in Figure 1. There is a substantial shift to the right in the distribution indicating a marked increase in the height of children, of a given age, over this period. This is also apparent in summary statistics, which we present in Table 1 for both the full and estimation samples. In each year, there is little difference between the samples, confirming that, after weighting, the restrictions we impose introduce little or no bias. For the full samples, median z-score increases from -2.07 in 1993 to -1.70 in 1998, an increase of almost one-fifth. The percentage of children who are stunted falls from 52% to 37%. Height increases at all percentiles, with the absolute changes being greater at lower percentiles but the relative changes being much larger at higher percentiles. As a result, the dispersion in height deficit increases, as is indicated by a 20% rise in the magnitude of the coefficient of variation. Not only does total inequality in child height deficit increase, height disparities by income also rise. This is indicated by a rise of one-fifth in the correlation coefficient between height-for-age z-scores and household consumption per capita. Given that income inequality increased during this period (see below), this rise in the income-child height gradient implies an even greater increase in the magnitude of income-related inequality in child height.

TABLE 1

Mean values of the covariates used in the quantile regressions are presented in Table 2 for both years, and both the full and estimation sample. In general, the restrictions imposed to arrive at the estimation sample do not affect the means, with slight exceptions for proportion of children with mothers and fathers absent from the household, belonging to ethnic minorities and living in some regions. The set of covariates includes child and household level characteristics—log of household consumption per capita, household hygiene conditions, child demographics, parental age, height and schooling, household size and age structure, ethnicity, urban dwelling and region—that have been found to explain variation in child height (Behrman and Deolalikar 1988; Thomas et al. 1990; Strauss and Thomas 1995; Glewwe et al. 2003). In a structural household production model of child nutrition (Behrman and Deolalikar 1988; Strauss and Thomas 1995), household consumption, sanitation and safe drinking water supply would all be considered endogenous since they are chosen conditional on the child's unobservable health stock that also determines nutritional status. The number of children in the household would also be endogenous in a model in which parents trade the quality against the quantity of children (Becker and Tomes 1976). Since our objective is not to identify the causal effects of these factors but simply to explain variation in child height, we need not worry about potential endogeneity. This does, however, restrict the interpretation that can be placed on the regression coefficients.

There are significant changes in the means of most of the covariates. Exceptions are the proportions that are male, have an absent or deceased mother or father, an ethnic minority, an urban dweller and located in certain regions. The observed increase in the mean of the log of consumption is equivalent to a 30% rise in real per capita consumption between 1993 and 1998. This substantial rise in average real consumption was accompanied by a rise in inequality, with the Gini coefficient for this sample rising from 0.31 to 0.34.⁶ Child nutrition status is not determined solely by nutritional intake. Disease and illness, particularly diarrhoea, draw on energy stocks and interfere with the absorption of nutrients (World Bank 2006). Sanitation and the quality of drinking water are therefore potentially important determinants of nutritional status and we include indicators of both. The percentage of

children living in a household with safe drinking water increased by almost 5 points over the period. The percentage with a sanitary toilet increased by almost 6 points.

Child malnutrition tends to vary with age. On average, it increases during the first two years of life and thereafter fluctuates with no clear trend (World Bank 2006). In the regressions, we include a full set of age dummies. In the Table 2, we present mean ages. There is an increase of around 6 months in the average age of children in the samples over the five-year period. This reflects the falling rate of fertility over the period (World Bank 2001), resulting in cohorts of older children being larger than those of younger children. Note that there is also a fall in the fraction of household members that are less than 6 years old and a rise in the fraction between 6 and 15 years. These differences may also be due to the panel nature of the survey. Even though new households were added to the second wave, it is likely that it contains fewer newly established households with very young children. Possibly for the same reason, there is an increase of around 7 months in the mean age of mothers.

There is an increase in the average height of both fathers and mothers of around 0.4 cm over the five-year period, which may be due to trend increase in the height of the population, and would be expected to make a positive contribution to the observed increase in the height of children. The same should be true of the observed slight increase in the years of schooling of both fathers and mothers.

TABLE 2

Before turning to the decomposition analysis, we present, in Table 3, estimates of conditional median regressions in order to show how child height is related to the covariates at the centre of the distribution in each year. In the decomposition analysis, these relationships are estimated across the full range of the conditional height-for-age distribution. In addition to the log of household consumption per capita, we also include in the regressions the log of the commune mean of per capita consumption. This is done to pick up the effects of commune level characteristics, such as public hygiene conditions and the availability and quality of health care, which have been found to impact on child nutritional status (Thomas and Strauss 1993; Thomas et al. 1996). Using commune dummies for this purpose is not computationally feasible given that there are up to 196 communes and 4500 quantile regressions are estimated for each year in the decomposition analysis. Data on commune level characteristics,

including health care facilities, are available only for the rural communes. The mean consumption in an area should act as a proxy for general living conditions. It may also have an impact on those conditions by determining resources available for investment in infrastructure and health care. In Vietnam, a share, admittedly small, of the financing of public health care is raised at the commune level. The results indicate that the (conditional) median child height-for-age z-score is rising significantly with both household and commune mean consumption. In fact, the relationship with commune mean consumption is stronger in both years. Perhaps surprisingly, household access to safe drinking water is not significantly correlated with median child height in either year and existence of a sanitary toilet within the household has a significant positive coefficient only in 1993.

Median nutritional status, as measured by the height-for-age z-score, is significantly lower for boys than for girls. This is a standard result (World Bank 2006). In both years, all age dummies are significantly negative relative to the reference group of infants less than 1 year old. This is consistent with the international evidence (World Bank 2006) that shows height deficits are usually smallest at birth and increase rapidly in the first two years of life before levelling off. The median z-score is increasing with the mother's age and the height of both parents. There is a significant positive relationship with mother's schooling only in 1993 and no significant relationship with father's schooling in either year. Median height deficit is greater in larger households and in households with a larger proportion of infants and children. In 1993, children from ethnic minorities were significantly shorter and those in an urban location significantly taller. But neither of these relationships remained significant in 1998. There are significant regional variations in both years.

TABLE 3

5 Results

5.1 Tests of parameter homogeneity and stability

As explained in section 3, quantile regression estimates of conditional distribution functions of height-for-age z-scores are used in combination with sampled sets of covariates to simulate marginal empirical and counterfactual distributions. The

advantage of quantile regression for this purpose is that it can give a more complete representation of the conditional distribution, allowing parameters to vary across the range of the distribution. A marginal distribution could also be simulated by using ordinary least squares (OLS) estimates of the conditional mean function to score a prediction and then adding a constant-variance random term. But the quantile procedure will be superior provided the conditional quantile functions differ from the conditional mean function in more than the intercept. We test the null of equality between all quantile regression slope parameters and respective OLS parameters for each of 19 evenly spaced quantiles from 0.05 to 0.95. The results of these Wald tests (Koenker and Bassett 1982) are reported in Table 4. Covariance matrices were obtained from 100 bootstrap estimates of the 19 conditional quantile functions and the conditional mean function. At the 1% significance level, we reject equality of the conditional quantile and mean functions for all but one quantile in 1993 and three in 1998. The four quantiles at which the null is not rejected are all relatively close to the conditional median, indicating less difference from the conditional mean function in this range. Away from the centre of the distribution, the null is always very decisively rejected. These results indicate failure of the normal location model for the conditional distribution of height-for-age z-scores and hence the superiority of quantile regression over OLS (Koenker 2005).

TABLE 4

Our objective is to decompose the change in the distribution of child height into that part due to the change in the distribution of covariates and that due to change in the relationships between height and covariates. In the previous section we documented substantial changes in the means of a number of covariates. We now test for significant changes in the parameters of the child height functions. Specifically, for each of the 19 conditional quantile functions used in the parameter homogeneity tests, we test the null that all parameters are stable between the two surveys. To implement this, we pool the data across the waves, estimate models with a full set of time interactions and test the joint significance of these interactions and a time dummy. Covariance matrices are constructed from 100 bootstrap replications for each quantile. The results, presented in Table 5, indicate that stability of the functions is very strongly rejected at all quantiles. This is not simply due to shifts in the intercepts

over time. We report tests for the stability of sets of parameters across all 19 of the conditional quantile functions. Stability is decisively rejected for the parameters on the age dummies, parental characteristics, household size and demographics, and region dummies. But stability of the household and commune mean consumption parameters at every quantile is not rejected. Neither is stability rejected for the parameters on safe drinking water and sanitation.

TABLE 5

5.2 Changes in child height explained by changes in all covariates

Figure 2 shows the simulated marginal distributions of height-for-age z-scores in both 1993 and 1998, along with a counterfactual density showing how child height would have been distributed if covariates were distributed as in 1998 but their relationships with child height remained as they were in 1993. Comparing the simulated 1993 distribution with the counterfactual identifies the contribution of changes in the covariates, while comparison between the simulated 1998 distribution and the counterfactual identifies the contribution of changes in the conditional quantile functions. Changes in the covariates appear to have contributed little to explanation of the decrease in very severe levels of malnutrition—z-scores below -4.5 . Above this point, the contribution of covariates increases and they explain most of the shift in the distribution just below the mode. Covariates appear to explain a greater share of shift in the top part than in the bottom part of the distribution, although in the top part there is still a sizable share that is explained by the changes in the conditional distribution.

In Table 6 we present results of the decomposition given in equation (1) for various parameters of the child height-for-age z-score distribution. In the first three columns, we present the estimates for each year and for the change between them. 95% confidence intervals are given in parentheses. These are calculated from the percentiles of 1000 bootstrap estimates. In the final three columns, we give the estimates of the change in the statistic explained by: (i) changes in the distributions of all covariates; (ii) changes in the coefficients of the conditional quantile functions; and, (iii) a residual term that is due to the differences between the empirical and simulated densities. Besides 95% confidence intervals for these contributions, we also express each as a percentage of the actual change in the statistic. These percentage contributions are also given using the reverse weighting, i.e. calculating the

covariates' contributions at 1998, rather than 1993, values of the coefficients and coefficients' contributions at 1993, rather than 1998, values of the covariates.

The median z-score is estimated to have risen by 0.35, with an increase of 0.18 (52%) explained by changes in the covariates and 0.22 (62%) explained by changes in the coefficients. Both contributions are significantly different from zero. These results suggest that if the distribution of child height conditional on covariates had remained as it was in 1993 and all that happened until 1998 was that covariates changed as observed, then one would have predicted an increase of 0.18 in the median z-score, just more than half of the actual increase. The other half or more of the increase in the median is explained by a shift in the distribution of child height conditional on the covariates. Changes in covariates also explain around one-half of the 14-percentage point decrease in the proportion of children who are stunted.

The relative contributions of changes in covariates and in coefficients vary across the range of the distribution, with covariates explaining relatively less of the decrease in severe degrees of malnutrition. For example, covariates explain only 31% of the increases in the 10th percentile of the height-for-age distribution but 45-50% of the increases in the top half of the distribution. This is the result of little systematic variation in the absolute contribution of covariates combined with greater absolute increases in z-scores in the bottom half of the distribution. The absolute contribution of changes in the parameters of the conditional quantile functions is largest at the bottom of the distribution and declines up to the 75th percentile before increasing slightly. This suggests that changes in conditions that produce gains in child nutrition for given values of the observable characteristics included in regressions and/or which change the impact of characteristics on child nutrition are more important in explaining the reduction in severe levels of malnourishment.

Changes in covariates and in coefficients explain 44% and 68% respectively of the increased disparity in child height deficit, measured by the coefficient of variation. The contribution of covariates to the increased income gradient, measured by the correlation coefficient with household per capita consumption, is much greater than it is for the other statistics. In fact, from the change in covariates alone, the correlation coefficient would have been anticipated to rise by 26% more than it actually did.

With the exception of only one statistic (the 75th percentile), the estimated contribution of changes in covariates tends to increase and there is a corresponding

decrease in the contribution of coefficients when the reverse weighting scheme is used for the decomposition. As a result, the estimates move in the direction of greater equality between the respective contributions. For the median and coefficient of variation, the point estimates of the covariates' contributions actually become greater than those of coefficients. Decomposition of the change in the proportion stunted is less sensitive to the weighting adopted.

TABLE 6

5.3 Changes in child height explained by income

In section 3, we described two methods—represented by equations (2) and (3)—of estimating the change in height explained by a particular covariate. Equation (2) gives the effect of changing the distribution of a covariate and all of its correlated determinants of child height. Equation (3) exploits the panel nature of the data to estimate the effect of changing one covariate while holding the others constant. We label these total and partial contributions respectively.

In Table 7, we present the estimates of the total contribution of change in the distribution of household consumption per capita and the partial contributions of variables or groups of variables, again focussing on consumption and related factors. As explained in section 3, the total contribution of consumption is estimated by re-sampling observations from the 1993 simulated density such that twenty evenly spaced percentile points of household consumption per capita correspond to those observed in 1998. Given that consumption increased over the period, we effectively over sample children from high consumption households and under sample those from poor households. In doing so, the distribution of all covariates that are correlated with household consumption will be altered and the estimated contribution to the change in child height will be that of changes in household consumption and all its observable correlates. Since household consumption is strongly correlated with commune mean consumption, the procedure will largely capture the aggregate contribution of both. The estimate of the change in the median z-score explained by the change in the distribution of consumption is 0.14, which is 41% of the actual change. This is very near as much as the estimated contribution of all covariates—0.18 (Table 6). One reason for this is that changes in covariates can have offsetting effects on child height. But it is also because the method simulates changes in some

covariates by degrees that are not actually observed. For example, since ethnicity is strongly correlated with consumption, the proportion of children from ethnic minorities in the counterfactual sample falls well below that observed in 1998. This is the result of simulating the effect of change in consumption by making poor children resemble better-off children.

By this method, the change in the distribution of household consumption explains 39% of the fall in the proportion of stunted children. The relative contribution of consumption to the increase in the 10th percentile point of the height distribution is much smaller (24%) than it is to the increases in the top quarter of the distribution (45-50%). This is not only because the magnitude of the z-score increase is largest at the 10th percentile but also because the change in consumption makes an absolutely smaller contribution at this percentile. The latter indicates that the 1993 quantile regression coefficients on household consumption and/or its correlates are smaller at lower quantiles and so predict smaller shifts in child height from changes in covariates.⁷

To estimate the partial contribution of covariates, we take observations used to compute the 1993 simulated height-for-age density, change 1993 values of one or more covariates to the values observed for the child's household in 1998 and keep other covariates at 1993 values. In consecutive columns, we first show the effect of changing household consumption only, then both household and commune mean consumption, then the dummies for drinking water and sanitation quality only, and then all the consumption and hygiene variables together. With the latter, we attempt to estimate the contribution of economic growth to the explanation of changes in height-for-age distribution through all of the observable factors that could plausibly respond to growth. From the change in the distribution of household consumption alone, we predict an increase in the median z-score of 0.07, which is just over one-third of the total contribution of consumption and 14% of the actual change in the median. If we allow both household and commune mean consumption to change, the predicted change in the median z-score doubles, accounting for one-third of the actual change. This reflects the strong positive coefficient on commune mean consumption, which can be observed in Table 3 but also exists for other conditional quantile functions. The increase in the proportion of children with access to safe drinking water and a sanitary toilet makes a much smaller contribution, explaining 3.8% of the increase in the median. Changing household consumption, commune mean consumption and the

hygiene indicators together, we predict a rise in the median of 0.13, or 38% of the actual increase. Hence, as would be expected, the more correlated factors we allow to change simultaneously, the closer we get to the estimated total contribution of consumption. But by the partial approach, we avoid simulating the effect of changes in factors that cannot plausibly be argued to change with income, such as ethnicity.

Decreases in household size and the proportion of young children in the household explain 17% of the actual increase in the median. This is consistent with a scenario in which as households become richer, they are choosing to have fewer children and to invest more in the human capital, including the nutritional status, of a smaller number of offspring. In this case, this household size and structure effect would also, in part, be attributable to economic growth.

TABLE 7

The pattern of the partial contribution results for the proportion of children stunted is similar to that for the median. Changes in household consumption, commune mean consumption, and water and sanitation quality jointly explain 35% of the decrease in the proportion stunted. The contribution of these three factors to the explanation of changes in height increases, in both absolute and relative terms, in moving from the bottom to the top of the distribution. Income-related factors have contributed least to the explanation of the increase in height of the most malnourished children. While the contribution of household consumption alone varies little across the distribution, the combined contribution of household and commune mean consumption is rising, suggesting that it is commune mean consumption that contributes more at higher percentiles. This is confirmed by the fact that the regression coefficients on commune mean consumption are larger at higher quantiles of the 1993 conditional height distribution. The contribution of household size and structure is also monotonically increasing as one moves up the height distribution. Income-related factors explain 41% of the increase in the coefficient of variation and 48% of the increase in the income gradient in nutritional status.

6 Conclusion

Between 1993 and 1998, Vietnam experienced a dramatic rightward shift and increase in dispersion in the distribution of child nutritional status that coincided with a period of economic boom but also economic reform, nutrition policy initiatives and health sector changes. We set out to explain the change in the distribution and, in particular, to determine how much of the changes in various parameters of the child height distribution can be explained by changes in the distribution of income. Our results show that one-half of the 15-point fall in the proportion of children chronically malnourished, or stunted, can be explain by changes in the distributions of covariates. The other half of the decline in malnutrition is explained by changes in the conditional distribution of child height. That is, change in the relationships between height and observable correlates, such as income, and/or shifts in the distribution conditional on observables.

We find that changes in the distribution of household per capita consumption—our measure of permanent income—can explain, in total, almost 40% of both the fall in the proportion of children stunted and the 20% increase in the mean height-for-age z-score. This is less than the 60% of the increase in the mean that (Edmonds 2004) attributes to income but more than the 20% of the 25% fall in the proportion of children less 5 years old who are stunted attributed to income growth by (Glewwe et al. 2003). The inconsistencies between these estimates are most likely due to differences in the extent to which other factors are held constant when simulating the contribution of income. Glewwe, Koch and Nguyen take a multivariate approach and hold other observable factors constant. Edmonds does not. We control for other factors in estimation of the income coefficients but simulate the total contribution of income by allowing the distribution of income and all of its correlated observables to change. Since this induces changes in covariates that cannot plausibly be argued to vary with income, we also simulate the partial contribution of income by holding covariates constant. By this method, household income explains only around 15% of the rise in the mean and the fall in the proportion stunted, which is much closer to the estimate of Glewwe, Koch and Nguyen. If we allow not only household income to change, but also the mean income within the household's commune and access to safe drinking water and sanitation, then 35% of the fall in the proportion of children stunted can be explained. The contribution through commune mean income is

particularly large, accounting for roughly half of the gains in child nutritional status that is explained by income growth. This suggests rising prosperity of a community may bring significant gains in nutrition through improvements in commune infrastructure and reduced negative public health externalities. Explanations of changes occurring across the range of the nutrition distribution differ. Changes in covariates explain about half of the shift in the top half of the height-for-age distribution but only just over 30% of the increase in the 10th percentile point. Growth in household consumption and its correlates explain 40-50% of the improvements in nutrition in the top part of the distribution but less than 25% of the increase in the 10th percentile. So, changes observable characteristics, including income, account for less of the reductions in very severe degrees of malnutrition than they do of reductions in milder malnutrition. This is partly due to the fact some covariates, in particular commune mean consumption and household size and age structure, are more strongly correlated with child height at higher points in the distribution. From given changes in these covariates, predicted changes in child height are larger at higher quantiles. But more important is the fact that the conditional distribution has shifted more at lower quantiles than it has at higher quantiles. Changes in the 'technology' of child nutrition production and the supply of unobservable inputs appear to have contributed most to the reduction in very severe degrees of malnutrition. Some of these changes may derive from nutrition and health policy initiatives introduced or extended in Vietnam during the 1990s. Examples include improvements in the energy and protein content of the diet, encouragement of breast feeding, salt iodisation, increased child immunisation, increased availability of oral rehydration therapy, reduced real price of drugs and medical care, and increased health insurance coverage of school children (World Bank 2001; Hop 2003). While the current analysis is far from an evaluation of the impact of these changes on child nutrition, the results are consistent with a scenario in which they have shifted the conditional quantile functions most at very low levels of nutrition. This does not seem implausible, particularly for those interventions for which correction of nutrition deficiencies was the primary objective.

The analysis covered children less than 10 years old. Restricting attention to children less than 5 years is also interesting since one is then comparing children exposed to conditions up to 1993 with those exposed only to conditions between 1993 and 1998. We have repeated the analysis for the younger age group. Qualitatively the results are similar to those reported here, with the quantitative difference being that

covariates explain even less of the changes and this contribution rises even more steeply in moving up the distribution.⁸ For example, covariates explain 45% of the increase in the median z-score, as opposed to 52% for the under tens, and only 10% of the increase in the 10th percentile point compared with 30% for the under tens. Changes in covariates explain only 35% of the fall in the proportion stunted compared with 50% for the under tens. Changes in conditions influencing the production of child nutritional status conditional on observable characteristics appear to have been even more important in explaining the gains in height of very young children. This is consistent with the fact that nutrition intervention programmes tend to focus on younger children.

Changes in observable factors explain one-half of the 15-point fall in the rate of chronic malnourishment of children under 10 years that occurred in Vietnam in the nineties. The other half of the gain could not have been predicted from the conditional distribution of child height that held in 1993. Forecasts made at that time would have been overly conservative. Of course, forecasts are always prone to error and the case of Vietnam may be particularly atypical. It is a country in transition and has undergone an unusually large number of reforms in recent years that may be beneficial for the production of child nutrition for any given change in income and its correlates. This said, as Edmonds (2004) has observed, the inability of income growth to explain all of the gains in child nutritional status has important implications for the definition and monitoring of development targets. Provided weight is given to non-economic dimensions of development, such as child nutrition, then it seems necessary to set explicit targets for them and to monitor them separately. The positive message from the Vietnam experience is that the gains in human development indicators can be greater than those predicted from trends in economic growth.

NOTES

¹ A height-for-age z-score is the difference between the height of a child and the median height of a child of the same age and sex in a well-nourished reference population divided by the standard deviation in the reference population.

² A child is defined as stunted if her height is more than two standard deviations below the median height in a well-nourished reference population, i.e. a height-for-age z-score less than -2 .

³ See (Nguyen et al. 2006) for an application of this decomposition technique to urban-rural inequality.

⁴ Moving from a situation in which oral rehydration salts are never available to one in which they are always available is associated with a 0.44 point rise in the height-for-age z-score, which is close to the actual change in the mean score in rural areas. But data on the change in the supply of rehydration salts over time are not available. It is only known that in 1998 they were available in 83% of commune health centres all of the time.

⁵ Because of the prohibitive computational cost, we do not include computation of the quantile regression coefficients within the bootstrap procedure. We estimate 4500 quantile regressions for each year to produce the conditional distributions. Bootstrapping these with 1000 replications, which is appropriate given we are interested in estimating parameters in the tails of the distribution, would require a total of 9 million quantile regressions.

⁶ These figures differ from those given in the introduction since they are derived from samples of households with children under the age of 10 years only.

⁷ Re-sampled covariates are combined randomly with the regression coefficients to produce the simulated and counterfactual distributions and so the magnitude of changes in covariates cannot explain differences across the range of the height-for-age distribution in the relative contribution of covariates.

⁸ Detailed results for the under 5 years age group are available from the authors on request.

References

- Becker, Gary S. and Nigel Tomes (1976). "Child endowments and the quantity and quality of children." Journal of Political Economy **84**(4): S143-S162.
- Behrman, J.R. and A.B. Deolalikar (1988). Health and Nutrition. Handbook of development economics. Volume 1. Handbooks in Economics. H. Chenery and T. N. E. Srinivasan. Amsterdam, New York and Oxford, Elsevier Science: 631-711.
- Blinder, Alan S. (1973). "Wage discrimination: reduced form and structural estimates." Journal of Human Resources **8**: 436-55.
- Edmonds, Eric (2004). How well do improvements in economic status track non-monetary measures of well-being? Evidence from child height. Dartmouth College. Hanover, NH.
- Glewwe, P (2003). An overview of economic growth and household welfare in Vietnam in the 1990s. Economic Growth, Poverty and Household Welfare: Policy Lessons from Vietnam. P. Glewwe, N. Agrawal and D. Dollar. Washington DC, World Bank.
- Glewwe, P, S Koch and BL Nguyen (2003). Childnutrition, economic growth, and the provision of health services in Vietnam. Economic Growth, Poverty and Household Welfare: Policy Lessons from Vietnam. P. Glewwe, N. Agrawal and D. Dollar. Washington DC, World Bank.
- Haddad, L., H. Alderman, S. Appleton, L. Song and Y. Yohannes (2003). "Reducing child malnutrition: How far does income growth take us?" World Bank Economic Review **17**(1): 107-131.
- Hop, Le Thi (2003). "Programs to Improve Production and Consumption of Animal Source Foods and Malnutrition in Vietnam." Journal of Nutrition **133**: 4006S-4009S.
- Koenker, Roger (2005). Quantile Regression. New York, Cambridge University Press.
- Koenker, Roger and Gib Bassett (1982). "Tests of linear hypotheses and l_1 estimation." Econometrica **50**(1577-1584).
- Machado, J.A.F. and J. Mata (2005). "Counterfactual decomposition of changes in wage distributions using quantile regression." Journal of Applied Econometrics **20**(4): 445-65.
- Nguyen, Binh T., James W. Albrecht, Susan B. Vroman and M. Daniel Westbrook (2006). "A quantile regression decomposition of urban-rural inequality in Vietnam." Journal of Development Economics **forthcoming**.
- Oaxaca, RL (1973). "Male-female wage differentials in urban labor markets." International Economic Review **14**: 693-709.
- Ponce, N, P Gertler and P Glewwe (1998). Will Vietnam grow out of malnutrition? Household Welfare and Vietnam's Transition. D. Dollar, P. Glewwe and J. Litvack. Washington DC, World Bank: 257-275.
- Smith, Lisa C and Laurence Haddad (2002). "How potent is economic growth in reducing undernutrition? What are the pathways of impact? New cross-country evidence." Economic Development and Cultural Change **51**(1): 55-76.
- Steckel, Richard H. (1995). "Stature and the Standard of Living." Journal of Economic Literature **33**(4): 1903-41.

- Strauss, J and D Thomas (1995). Human resources: Empirical modeling of household and family decisions. Handbook of Development Economics vol. III. J. Behrman and T. Srinivasan. Amsterdam, North Holland.
- Thomas, D., V. Lavy and J. Strauss (1996). "Public policy and anthropometric outcomes in Cote d'Ivoire." Journal of Public Economics **61**: 155-92.
- Thomas, Duncan and John Strauss (1993). "Prices, infrastructure, household characteristics and child height." Journal of Development Economics **39**: 301-31.
- Thomas, Duncan, John Strauss and Maria Helena Henriques (1990). "Child survival, height-for-age and household characteristics in Brazil." Journal of Development Economics **33**(2): 197-234.
- World Bank (2001). Vietnam Growing healthy: A review of Vietnam's health sector. Hanoi, The World Bank.
- World Bank (2006). Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action. Washington, D.C., The World Bank.
- World Health Organisation (1995). Physical status: Use and interpretation of anthropometry. Report of a WHO Expert Committee. WHO Technical Report 854. Geneva, WHO.

FIGURES

Figure 1:

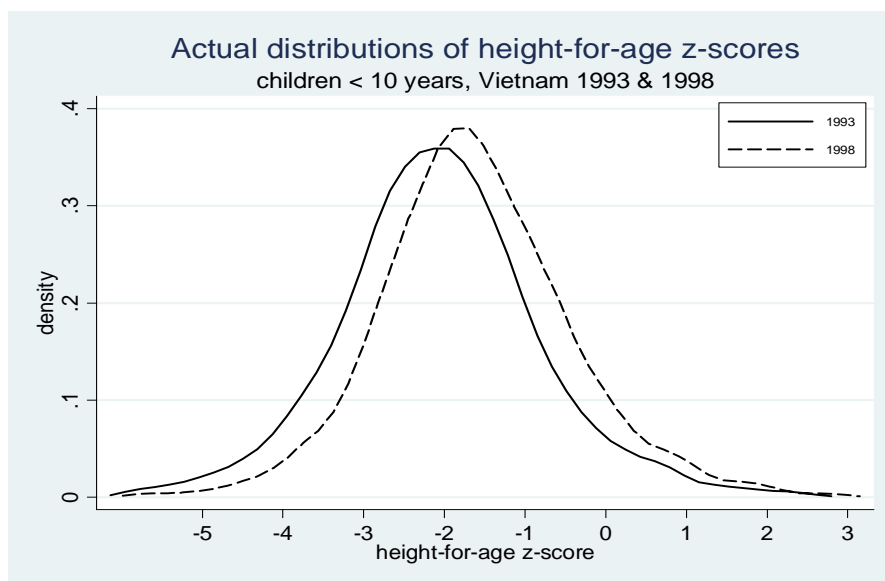


Figure 2:

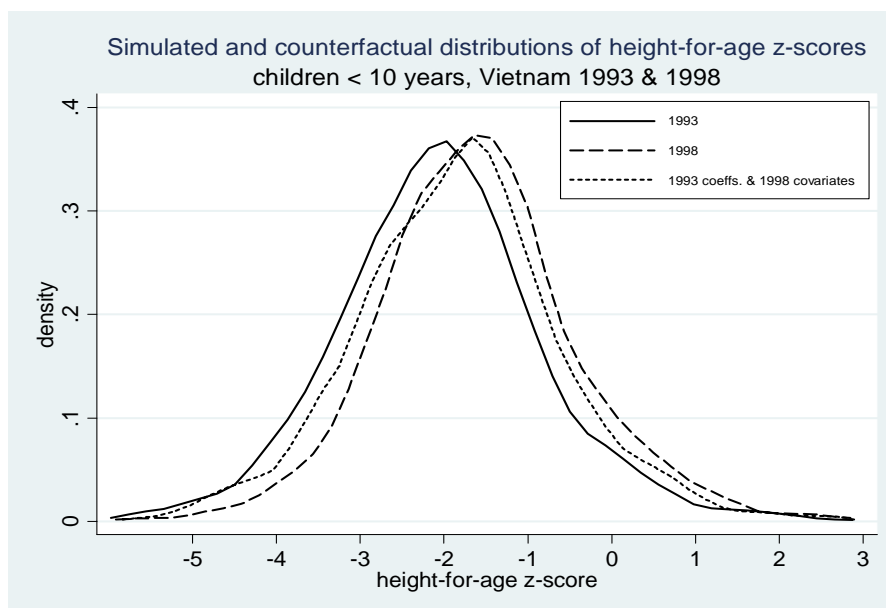


Table 1: Descriptive statistics for height-for-age z-scores, children less than 10 years old

| | Full sample ^a | | Estimation sample ^b | |
|---|--------------------------|----------|--------------------------------|----------|
| | 1993 | 1998 | 1993 | 1998 |
| Proportion stunted ^c | 0.5255 | 0.3745 | 0.5221 | 0.3782 |
| (standard error) ^d | (0.0123) | (0.0134) | (0.0130) | (0.0138) |
| 10th percentile | -3.52 | -3.01 | -3.52 | -3.02 |
| 25th percentile | -2.81 | -2.36 | -2.79 | -2.37 |
| Median | -2.07 | -1.70 | -2.06 | -1.71 |
| 75th percentile | -1.31 | -0.94 | -1.31 | -0.95 |
| 90th percentile | -0.50 | -0.15 | -0.50 | -0.17 |
| Mean | -2.0334 | -1.6248 | -2.0308 | -1.6359 |
| (standard error) | (0.0327) | (0.0341) | (0.0350) | (0.0355) |
| Coefficient of variation | -0.6142 | -0.7257 | -0.6109 | -0.7203 |
| Correlation coefficient with household consumption per capita | 0.2191 | 0.2650 | 0.2193 | 0.2771 |
| Number of observations | 5692 | 5193 | 4794 | 4852 |

Notes:

- a. All children less than 10 years old in respective year.
- b. Full sample less those with missing values on any covariates. 1993 sample excludes children in households that had attrited from survey by 1998.
- c. Weights applied in computation of all statistics.
- d. Standard errors corrected for sample stratification and clustering.

Table 2: Means of covariates and tests of no change in them over time (weights applied)

| | Full sample ^a | | | Estimation sample ^b | | |
|---|--------------------------|-------------------|----------------------|--------------------------------|--------|----------------------|
| | 1993 | 1998 | p-value ^c | 1993 | 1998 | p-value ^c |
| log household consumption per capita ^d | 7.2964 | 7.5605 | 0.000 | 7.3040 | 7.5600 | 0.000 |
| safe drinking water (dummy) ^e | 0.7805 | 0.8288 | 0.000 | 0.7829 | 0.8265 | 0.000 |
| sanitary toilet (dummy) ^f | 0.1240 | 0.1832 | 0.000 | 0.1276 | 0.1805 | 0.000 |
| male (dummy) | 0.5114 | 0.5048 | 0.496 | 0.5142 | 0.5051 | 0.383 |
| age | 5.0588 | 5.5774 | 0.000 | 5.0931 | 5.6057 | 0.000 |
| mother's age ^g | 32.63 | 33.22 | 0.000 | 32.64 | 33.19 | 0.000 |
| mother's height (cm) ^g | 151.79 | 152.28 | 0.000 | 151.79 | 152.27 | 0.000 |
| father's height (cm) ^g | 161.87 | 162.26 | 0.000 | 161.88 | 162.26 | 0.001 |
| mother's years of schooling ^h | 6.2502 | 6.6522 | 0.000 | 6.2602 | 6.6138 | 0.000 |
| father's years of schooling ^h | 7.2170 | 7.4442 | 0.000 | 7.2353 | 7.4341 | 0.004 |
| mother dead or absent (dummy) | 0.0243 | 0.0259 | 0.591 | 0.0227 | 0.0206 | 0.488 |
| father dead or absent (dummy) | 0.0855 | 0.0803 | 0.327 | 0.0793 | 0.0703 | 0.104 |
| household size | 6.0373 | 5.7605 | 0.000 | 6.0423 | 5.7803 | 0.000 |
| kids < 6 yrs. as proportion of household size | 0.2465 | 0.1961 | 0.000 | 0.2462 | 0.1954 | 0.000 |
| kids 6-15 yrs. as proportion of household size | 0.2748 | 0.2950 | 0.000 | 0.2765 | 0.2988 | 0.000 |
| ethnic minority (dummy) ⁱ | 0.1693 | 0.1831 | 0.066 | 0.1752 | 0.1881 | 0.107 |
| urban (dummy) | 0.1547 | 0.1590 | 0.518 | 0.1556 | 0.1580 | 0.750 |
| Red River delta (dummy) | 0.1980 | 0.1687 | 0.000 | 0.1964 | 0.1650 | 0.000 |
| Northeast (dummy) | 0.1613 | 0.1729 | 0.125 | 0.1672 | 0.1754 | 0.310 |
| Northwest (dummy) | 0.0348 | 0.0334 | 0.703 | 0.0365 | 0.0348 | 0.663 |
| North Central coast (dummy) | 0.1432 | 0.1673 | 0.001 | 0.1420 | 0.1659 | 0.002 |
| South Central coast (dummy) | 0.0829 | 0.0848 | 0.697 | 0.0784 | 0.0799 | 0.775 |
| Central Highlands (dummy) | 0.0295 | 0.0346 | 0.077 | 0.0282 | 0.0362 | 0.009 |
| Southeast (dummy) | 0.1501 | 0.1475 | 0.693 | 0.1560 | 0.1492 | 0.326 |
| Number of observations | 5864 ^j | 5324 ^j | | 4895 | 4943 | |

Notes:

- a. All children less than 10 years old in respective year.
- b. Full sample less those with missing values on any covariates. 1993 sample excludes children in households that had attrited from survey by 1998.
- c. p-value from test of equality of 1993 and 1998 means.
- d. In constant (January 1998) prices.
- e. 1 if household's main source of drinking water meets WHO and UNICEF (2004) definition of improved drinking water source.
- f. 1 if household has private flush toilet or latrine.
- g. Mean value given if parent is dead or absent.
- h. Not including post-high school. Data available even if parent dead or absent.
- i. 1 if does not belong to Kinh tribe.
- j. The number of observations is less than this for variables with missing values.

Table 3: Conditional median functions of height-for-age z-scores

| | 1993 | | | 1998 | | |
|---|-------------|---------------------------|---------|-------------|---------------------------|---------|
| | Coefficient | stand. error ^a | p-value | Coefficient | stand. error ^a | p-value |
| log household consumption per capita | 0.1394 | 0.0583 | 0.0170 | 0.1630 | 0.0401 | 0.0000 |
| log commune mean household consumption per capita | 0.2348 | 0.0822 | 0.0040 | 0.3176 | 0.0596 | 0.0000 |
| safe drinking water | 0.0206 | 0.0601 | 0.7320 | 0.0128 | 0.0509 | 0.8010 |
| sanitary toilet | 0.1391 | 0.0551 | 0.0120 | 0.0188 | 0.0542 | 0.7290 |
| male | -0.1090 | 0.0340 | 0.0010 | -0.1245 | 0.0289 | 0.0000 |
| age 1-2 years | -1.6895 | 0.1162 | 0.0000 | -1.3063 | 0.1049 | 0.0000 |
| age 2-3 years | -1.4316 | 0.1125 | 0.0000 | -1.0546 | 0.0898 | 0.0000 |
| age 3-4 years | -1.4594 | 0.1059 | 0.0000 | -1.2870 | 0.0905 | 0.0000 |
| age 4-5 years | -1.5806 | 0.1066 | 0.0000 | -1.4431 | 0.0891 | 0.0000 |
| age 5-6 years | -1.5172 | 0.1134 | 0.0000 | -1.3957 | 0.0834 | 0.0000 |
| age 6-7 years | -1.5330 | 0.1092 | 0.0000 | -1.4129 | 0.0912 | 0.0000 |
| age 7-8 years | -1.4782 | 0.1085 | 0.0000 | -1.3061 | 0.1007 | 0.0000 |
| age 8-9 years | -1.3933 | 0.1080 | 0.0000 | -1.3041 | 0.0987 | 0.0000 |
| age 9-10 years | -1.4849 | 0.1128 | 0.0000 | -1.3914 | 0.0952 | 0.0000 |
| mother's age | 0.0136 | 0.0040 | 0.0010 | 0.0085 | 0.0027 | 0.0020 |
| mother's height (cm) | 0.0427 | 0.0040 | 0.0000 | 0.0485 | 0.0036 | 0.0000 |
| father's height (cm) | 0.0417 | 0.0033 | 0.0000 | 0.0447 | 0.0035 | 0.0000 |
| mother's years of schooling | 0.0144 | 0.0071 | 0.0440 | 0.0072 | 0.0066 | 0.2750 |
| father's years of schooling | 0.0066 | 0.0069 | 0.3390 | 0.0035 | 0.0062 | 0.5750 |
| mother dead or absent | -0.1883 | 0.1731 | 0.2770 | -0.0803 | 0.0920 | 0.3830 |
| father dead or absent | -0.0396 | 0.0667 | 0.5520 | -0.0469 | 0.0616 | 0.4470 |
| household size | -0.0276 | 0.0091 | 0.0020 | -0.0291 | 0.0085 | 0.0010 |
| kids < 6 yrs. as proportion of household size | -0.6710 | 0.1718 | 0.0000 | -0.8043 | 0.1735 | 0.0000 |
| kids 6-15 yrs. as proportion of household size | -0.5154 | 0.1549 | 0.0010 | -0.7285 | 0.1332 | 0.0000 |
| ethnic minority | -0.1548 | 0.0536 | 0.0040 | -0.0325 | 0.0427 | 0.4470 |
| urban | 0.1125 | 0.0598 | 0.0600 | 0.0588 | 0.0666 | 0.3780 |
| Red River delta | -0.2205 | 0.0814 | 0.0070 | -0.1199 | 0.0650 | 0.0650 |
| Northeast | -0.2577 | 0.0852 | 0.0020 | -0.1262 | 0.0614 | 0.0400 |
| Northwest | -0.0819 | 0.1136 | 0.4710 | 0.0717 | 0.1082 | 0.5080 |
| North Central coast | -0.3760 | 0.0871 | 0.0000 | -0.1920 | 0.0636 | 0.0030 |
| South Central coast | -0.1447 | 0.0795 | 0.0690 | -0.1829 | 0.0760 | 0.0160 |
| Central Highlands | -0.1220 | 0.1487 | 0.4120 | -0.0929 | 0.0820 | 0.2570 |
| Southeast | 0.0906 | 0.0771 | 0.2400 | 0.0731 | 0.0585 | 0.2120 |
| intercept | -16.6073 | 0.9031 | 0.0000 | -18.4439 | 0.9334 | 0.0000 |
| Pseudo R ² | 0.1403 | | | 0.1698 | | |
| Number of observations | 4794 | | | 4852 | | |

Notes:

a. Bootstrap standard error with 100 replications.

Table 4: Wald tests of equality of between slope parameters of each conditional quantile function and the conditional mean function

| Quantile | 1993 | | 1998 | |
|----------|------------------|---------|------------------|---------|
| | test statistic | p-value | test statistic | p-value |
| | $\sim\chi^2(33)$ | | $\sim\chi^2(33)$ | |
| 0.05 | 283.77 | 0.0000 | 139.84 | 0.0000 |
| 0.10 | 223.04 | 0.0000 | 149.66 | 0.0000 |
| 0.15 | 145.76 | 0.0000 | 126.26 | 0.0000 |
| 0.20 | 229.05 | 0.0000 | 175.33 | 0.0000 |
| 0.25 | 170.10 | 0.0000 | 170.44 | 0.0000 |
| 0.30 | 126.41 | 0.0000 | 130.42 | 0.0000 |
| 0.35 | 69.73 | 0.0002 | 119.47 | 0.0000 |
| 0.40 | 65.08 | 0.0007 | 78.53 | 0.0000 |
| 0.45 | 58.96 | 0.0036 | 52.87 | 0.0155 |
| 0.50 | 55.44 | 0.0086 | 67.38 | 0.0004 |
| 0.55 | 56.12 | 0.0073 | 61.63 | 0.0018 |
| 0.60 | 32.01 | 0.5161 | 45.03 | 0.0790 |
| 0.65 | 77.82 | 0.0000 | 43.05 | 0.1132 |
| 0.70 | 120.69 | 0.0000 | 63.96 | 0.0010 |
| 0.75 | 149.44 | 0.0000 | 78.02 | 0.0000 |
| 0.80 | 115.16 | 0.0000 | 74.28 | 0.0001 |
| 0.85 | 119.74 | 0.0000 | 144.46 | 0.0000 |
| 0.90 | 214.84 | 0.0000 | 173.27 | 0.0000 |
| 0.95 | 267.35 | 0.0000 | 146.58 | 0.0000 |

Notes:

At each quantile, the null tested is that each of 33 slope parameters is equal to the respective parameter of the conditional mean function estimated by OLS.

Covariance matrices constructed from 100 bootstrap replications.

Table 5: Wald tests of stability of conditional quantile functions between 1993 and 1998

| Stability of each conditional quantile function ^a | | | Stability of all quantile parameters for groups of covariates ^b | | | |
|--|------------------------------------|---------|--|------------------------------------|-------------------------|---------|
| Quantile | test statistic $\sim\chi^2(34)$ | p-value | Covariate | test statistic $\sim\chi^2(df)$ | degrees of freedom (df) | p-value |
| 0.05 | 140.92 | 0.0000 | household & commune mean consumption per capita | 41.3 | 38 | 0.3266 |
| 0.10 | 215.46 | 0.0000 | safe drinking water & sanitary toilet | 39.7 | 38 | 0.3956 |
| 0.15 | 208.06 | 0.0000 | age dummies | 1150.6 | 171 | 0.0000 |
| 0.20 | 190.99 | 0.0000 | parents' characteristics ^c | 467.5 | 133 | 0.0000 |
| 0.25 | 175.06 | 0.0000 | household size & demographics ^d | 420.6 | 76 | 0.0000 |
| 0.30 | 177.35 | 0.0000 | region dummies | 2032.6 | 133 | 0.0000 |
| 0.35 | 194.37 | 0.0000 | | | | |
| 0.40 | 205.38 | 0.0000 | | | | |
| 0.45 | 188.71 | 0.0000 | | | | |
| 0.50 | 161.64 | 0.0000 | | | | |
| 0.55 | 120.82 | 0.0000 | | | | |
| 0.60 | 132.09 | 0.0000 | | | | |
| 0.65 | 152.11 | 0.0000 | | | | |
| 0.70 | 112.24 | 0.0000 | | | | |
| 0.75 | 120.75 | 0.0000 | | | | |
| 0.80 | 99.84 | 0.0000 | | | | |
| 0.85 | 85.69 | 0.0000 | | | | |
| 0.90 | 78.39 | 0.0000 | | | | |
| 0.95 | 63.59 | 0.0015 | | | | |

Notes:

- At each quantile, the null is that the intercept and each of the 33 slope parameters are all constant between 1993 and 1998.
 - For each set of covariates, the null is that at every quantile the parameters on these covariates are constant between 1993 and 1998.
 - Mother's age, mother's and father's height and years of schooling.
 - Household size and age composition, and ethnicity.
- Covariance matrices are constructed from 100 bootstrap replications.

Table 6: Decomposition of changes in parameters of the distribution of height-for-age z-scores into the contribution of all covariates and all coefficients

| | Sample estimates ^a | | | Decomposition of changes ^b | | |
|---|-------------------------------|--------------------|--------------------|---------------------------------------|--------------------|-------------------|
| | 1993 | 1998 | change | Contributions to change of: | | |
| | | | | covariates | coefficients | residual |
| Proportion stunted | 0.5114 | 0.3715 | -0.1399 | -0.0693 | -0.0745 | 0.0040 |
| 95% confidence interval ^c | (0.4821, 0.5413) | (0.3414, 0.4018) | (-0.1810, -0.0984) | (-0.0880, -0.0287) | (-0.0960, -0.0679) | (-0.0298, 0.0327) |
| contribution as % of change | | | | 49.6% | 53.3% | -2.8% |
| contribution as % of change with reverse weighting ^d | | | | 50.4% | 52.4% | |
| 10th percentile | -3.52 | -3.02 | 0.500 | 0.1572 | 0.3852 | -0.0425 |
| | (-3.67, -3.39) | (-3.15, -2.90) | (0.340, 0.705) | (0.0181, 0.2840) | (0.3021, 0.4370) | (-0.1544, 0.1294) |
| | | | | 31.4% | 77.0% | -8.5% |
| | | | | 39.3% | 69.2% | |
| 25th percentile | -2.79 | -2.37 | 0.420 | 0.1474 | 0.2965 | -0.0239 |
| | (-2.875, -2.705) | (-2.47, -2.28) | (0.290, 0.540) | (0.0535, 0.2527) | (0.2269, 0.3236) | (-0.1018, 0.0854) |
| | | | | 35.1% | 70.6% | -5.7% |
| | | | | 46.3% | 59.4% | |
| median | -2.06 | -1.71 | 0.350 | 0.1803 | 0.2168 | -0.0471 |
| | (-2.14, -1.98) | (-1.79, -1.63) | (0.245, 0.460) | (0.0808, 0.2377) | (0.1743, 0.2509) | (-0.1063, 0.0620) |
| | | | | 51.5% | 61.9% | -13.5% |
| | | | | 59.8% | 53.6% | |
| 75th percentile | -1.31 | -0.95 | 0.360 | 0.1633 | 0.1966 | 0.0001 |
| | (-1.40, -1.24) | (-1.04, -0.86) | (0.250, 0.495) | (0.0686, 0.2299) | (0.1690, 0.2573) | (-0.0975, 0.1112) |
| | | | | 45.4% | 54.6% | 0.0% |
| | | | | 39.3% | 60.7% | |
| 90th percentile | -0.50 | -0.17 | 0.330 | 0.1662 | 0.2336 | -0.0698 |
| | (-0.64, -0.39) | (-0.27, -0.03) | (0.180, 0.530) | (0.0194, 0.2225) | (0.1507, 0.2976) | (-0.1508, 0.1691) |
| | | | | 50.4% | 70.8% | -21.2% |
| | | | | 58.9% | 62.3% | |
| Mean | -2.0308 | -1.6359 | 0.3949 | 0.1645 | 0.2557 | -0.0252 |
| | (-2.1079, -1.9516) | (-1.7188, -1.5556) | (0.2831, 0.5059) | (0.0695, 0.2258) | (0.2323, 0.2774) | (-0.0832, 0.0632) |
| | | | | 41.6% | 64.7% | -6.4% |
| | | | | 46.5% | 59.8% | |
| Coefficient of variation | -0.6109 | -0.7203 | -0.1094 | -0.0485 | -0.0741 | 0.0132 |
| | (-0.6456, -0.5761) | (-0.7629, -0.6809) | (-0.1633, -0.0580) | (-0.070, -0.0091) | (-0.0906, -0.0500) | (-0.0490, 0.0479) |
| | | | | 44.3% | 67.8% | -12.1% |
| | | | | 58.3% | 53.8% | |
| Correlation coefficient with household consumption per capita | 0.2193 | 0.2771 | 0.0578 | 0.0729 | -0.0133 | -0.0018 |
| | (0.1694, 0.2703) | (0.2197, 0.3277) | (-0.0183, 0.1314) | (0.0083, 0.1322) | (-0.0487, 0.0209) | (-0.066, 0.0664) |
| | | | | 126.2% | -23.1% | -3.1% |
| | | | | 133.2% | -30.1% | |

Notes:

a. Sample weights are applied in estimation of all parameters.

b. In the first 3 rows for each statistics, the covariates' contribution is calculated at the 1993 values of coefficients and the coefficients' contribution at 1998 values of the covariates.

c. Interval is the 0.025 and 0.975 percentile values of 1000 bootstrap estimates. The stratified and cluster sample design is replicated in drawing the bootstrap samples.

d. Covariates' contribution calculated at 1998 values of coefficients and coefficients' contribution at 1993 values of covariates.

Table 7: Decomposition of changes in parameters of the distribution of height-for-age z-scores into the contribution of specific covariates

| | Change: 1998 value - 1993 value | Contributions to change | | | | | |
|---|---------------------------------------|--|------------------------------------|--|------------------------------------|---|---|
| | | Total contribution of household consumption ^a | hhold. consumption ^c | hhold. & commune consumption ^d | water & sanitation ^e | consumption, water & sanitation ^f | hhold. size & structure ^g |
| Proportion stunted | -0.1399 | -0.0542 | -0.0202 | -0.0439 | -0.0049 | -0.0493 | -0.0230 |
| 95% confidence interval ^c | | (-0.0816, -0.0346) | (-0.0314, -0.0062) | (-0.0552, -0.0293) | (-0.0159, 0.0074) | (-0.0597, -0.0338) | (-0.0341, -0.0091) |
| contribution as % of change | | 38.8% | 14.4% | 31.4% | 3.5% | 35.2% | 16.4% |
| 10th percentile | 0.5000 | 0.1200 | 0.0772 | 0.1032 | 0.0067 | 0.1050 | 0.0446 |
| | | (0.0846, 0.2716) | (0.0203, 0.1433) | (0.0372, 0.1686) | (-0.0491, 0.0685) | (0.0523, 0.1779) | (-0.0135, 0.1119) |
| | | 24.0% | 15.4% | 20.6% | 1.3% | 21.0% | 8.9% |
| 25th percentile | 0.4200 | 0.1749 | 0.0693 | 0.1082 | 0.0097 | 0.1216 | 0.0426 |
| | | (0.0929, 0.2361) | (0.0291, 0.1050) | (0.0723, 0.1536) | (-0.0282, 0.0470) | (0.0803, 0.1602) | (0.0105, 0.0862) |
| | | 41.6% | 16.5% | 25.8% | 2.3% | 29.0% | 10.1% |
| median | 0.3500 | 0.1416 | 0.0503 | 0.1167 | 0.0134 | 0.1321 | 0.0607 |
| | | (0.0920, 0.2242) | (0.0164, 0.0852) | (0.0811, 0.1534) | (-0.0208, 0.0453) | (0.0923, 0.1652) | (0.0253, 0.0936) |
| | | 40.5% | 14.4% | 33.3% | 3.8% | 37.8% | 17.3% |
| 75th percentile | 0.3600 | 0.1831 | 0.0535 | 0.1444 | 0.0052 | 0.1444 | 0.0914 |
| | | (0.0778, 0.2263) | (0.0053, 0.0856) | (0.0804, 0.1651) | (-0.0293, 0.0472) | (0.0888, 0.1732) | (0.0397, 0.1179) |
| | | 50.9% | 14.9% | 40.1% | 1.5% | 40.1% | 25.4% |
| 90th percentile | 0.3300 | 0.1531 | 0.0818 | 0.1539 | 0.0186 | 0.1719 | 0.0991 |
| | | (0.0264, 0.2307) | (-0.0120, 0.1333) | (0.0562, 0.2146) | (-0.0604, 0.0814) | (0.0753, 0.2264) | (0.0126, 0.1620) |
| | | 46.4% | 24.8% | 46.6% | 5.6% | 52.1% | 30.0% |
| Mean | 0.3949 | 0.1528 | 0.0624 | 0.1223 | 0.0089 | 0.1312 | 0.0650 |
| | | (0.0962, 0.2149) | (0.0461, 0.0754) | (0.1027, 0.1388) | (-0.0043, 0.0233) | (0.1098, 0.1498) | (0.0499, 0.0802) |
| | | 38.7% | 15.8% | 31.0% | 2.2% | 33.2% | 16.5% |
| Coefficient of variation | -0.1094 | -0.0408 | -0.0159 | -0.0422 | -0.0025 | -0.0453 | -0.0258 |
| | | (-0.0733, -0.0126) | (-0.0317, 0.0000) | (-0.0583, -0.0252) | (-0.0183, 0.0126) | (-0.0629, -0.0281) | (-0.0419, -0.0095) |
| | | 37.3% | 14.6% | 38.5% | 2.3% | 41.4% | 23.6% |
| Correlation coefficient with hhold. consumption p.c. | 0.0578 | 0.0241 | 0.0140 | 0.0227 | 0.0031 | 0.0280 | -0.0051 |
| | | (-0.0147, 0.0680) | (-0.0307, 0.0517) | (-0.0220, 0.0616) | (-0.0348, 0.0379) | (-0.0168, 0.0651) | (-0.0409, 0.0296) |
| | | 41.7% | 24.2% | 39.3% | 5.4% | 48.4% | -8.7% |

Notes:

a. Calculated as in equation (2).

b. Calculated as in equation (3).

c. Household consumption per capita set to 1998 value observed for child's household and all other covariates held at 1993 values.

d. Household consumption per capita set and commune mean consumption set to 1998 values observed for child's household and all other covariates held at 1993 values.

e. Dummies for safe water and sanitation to 1998 values observed for child's household and all other covariates held at 1993 values.

f. Household consumption, commune mean consumption and water/sanitation dummies set to 1998 values observed for child's household and all other covariates held at 1993 values.

g. Household size and age structure set to 1998 value observed for child's household and all other covariates held at 1993 values.